**METHODOLOGIES:**

**DataSet For Research:**

**ISIC Archive Dataset :**

The ISIC Archive dataset contains around 23,000 pictures of skin issues, especially melanoma, an unsafe kind of skin sickness. These photographs Sourced from a collaborative effort of medical professionals, researchers. They're extraordinary quality. Experts looked at each picture and said if it's melanoma or not. This dataset is truly helpful for planning computers to check for melanoma, and it's a big deal in skin threatening development research. Notwithstanding, using it will in general be fascinating considering the way that there aren't various melanoma cases, and we ought to look out.

**Challenges in Gathering and Utilizing ISIC Dataset:**

**Data Irregularity:** The ISIC dataset, in the same way as other clinical datasets, experiences information unevenness, as malignant cases are relatively rare compared to benign ones. At the point when you select just a few pictures, guaranteeing a balanced portrayal in your training and testing datasets can challenge.

**Data Choice:** Picking the right subset of pictures to address both harmless(bengin) and dangerous cases(maligant) can be interesting. One-sided or deficient example determination might prompt one-sided computer based intelligence models.

**Data Augumentation:**

We can Augment our training data if necessary to increase the diversity of our dataset and improve model generalization. (Augmentation means rectifying the data imbalance and to avoid underfitting).There are many methods of augmentation like flipping ,rotating. We can also use Generative Adversial Network(GAW) to perform data augumentation.

**Generative Adversial Network (GAN):**It is a neural network architecture with 2 parts namely, generator, discriminator. Generator creates realistic looking images, discriminator distingushes between real and generated images.

Combine the original training dataset and the synthetic melanoma images generated by the GAN to create an augmented dataset. **A neural network called the Generator "G" produces fictional data "G" from noise "z.”[22]**

[log(D(x))] - [log(1 - D(G(z)))],

represents the loss function used in GANs, where D(x) is the discriminator's output when it evaluates real data (images), and D(G(z)) is the discriminator's output when it evaluates data generated by the generator from random noise z.

**Color Enhancement:** Enhancement technique based on the blue component of the RGB color channels is selected for the rest of processing which gives better results of both hair removal and segmentation**[23]**. we can find and distinguish hairlines in dermoscopic images. We do this by using a method called the "2-D derivatives of Gaussian" or DOG, focusing on the blue part of the images. It's like a special tool that helps us see lines of hair in four different directions. After we do this, we set a specific limit to separate the hairlines from the background in the image. It's kind of like drawing a clear line to say, "This is where the hair is, and this is everything else.

**Convolutional Neural Networks(CNNS):** The reception of Convolutional Brain Organizations (CNNs) has been a progressive move toward skin disease expectations. CNNs have to be trained using a huge dataset, This problem can be solved using the data augmentation. **We have used Traditional CNN for Skin Cancer Detection as mentioned in[24]** .CNNs are great at capturing skin lesions' intricate spatial hierarchies in medical images. The adoption of Convolutional Neural Networks (CNNs) has been a groundbreaking step in skin disease predictions. CNNs, such as VGG16, ResNet, and DenseNet, are employed as feature extractors and classifiers, enabling the automated identification of malignant skin lesions. These networks use convolution and pooling layers to extract relevant features from images, facilitating accurate classification.

**Transfer learning:**Transfer learning is a crucial technique for predicting skin cancer. Researchers can harness the vast knowledge stored in pre-trained CNN models, such as VGG16, ResNet-50, or DenseNet, which have been trained on large datasets like ImageNet. These pre-trained models have already learned to recognize general features such as edges, shapes, and textures. To adapt these models for the specific task of skin cancer classification, they can be fine-tuned using skin lesion datasets. This process is particularly valuable when working with limited labeled data since it significantly reduces the volume of data needed for training.

**VGG16, ResNet-50, or DenseNet For Feature Extraction:**

**The purpose of feature extraction is to reduce the number of features from images of the dataset.[25]** When we Load a pre-trained model e.g., VGG16 or ResNet50 **[26]**which involves using a neural network architecture that has already learned rich features from a large dataset. These pre-trained models are typically comprised of convolutional layers for feature extraction and fully connected layers for classification. To utilize these models for our own classification task, we remove the top classification layers, keeping only the feature extraction part. Then, you pass your augmented training data through these layers to extract abstract features from your images. These features are then flattened or reshaped into a format suitable for input to classifiers like XGBoost or LightGBM.

**Train XGBoost or LightGBM For Classification:**

**The most common way of preparing a XGBoost or LightGBM classifier includes utilizing the features extricated from your increased dataset as information. These elements basically address the significant data gained from our pictures.** Various AI draws near, such as SVM, KNN, Naive Bayes and others, are frequently used to distinguish between innocuous and cutaneous malignant melanoma injury**.[22] We take these component vectors and feed them into the classifier, which figures out how to make expectations in light of examples in the information.**

During training, the classifier fine-tunes its inner workings to reduce the gap between its predictions and the actual labels of the skin conditions. This means it gets better at understanding the relationships within the data, making it a reliable tool for predictions. In the end, you have a well-trained classifier that can be used to predict whether a new, unseen skin condition falls into the category of melanoma or not. This whole process is essential for your melanoma classification task.

**We are combining the techniques from both CNNs (for feature extraction) from [26] and ensemble learning for classification as in [27].** This approach leverages the strengths of each method: CNNs for feature extraction and XGBoost or LightGBM for making predictions based on those features. It's a powerful approach, especially when you have limited data or computational resources for training a full end-to-end deep neural network.

**Result:**

Recent advances in deep learning have demonstrated promising accuracy levels for skin cancer classification, rivaling dermatologist performance on select benchmarks. Key techniques include data augmentation, transfer learning, and uncertainty quantification.

Daghrir et al. [23] developed a hybrid approach combining multiple models - a convolutional neural network (CNN), support vector machine (SVM), and k-nearest neighbours (KNN) classifier. Their dataset contained 640 dermoscopic images resized to 224x224 pixels, split into 512 training and 128 testing images. The CNN architecture consisted of 3 convolutional layers with ReLU activation and 3x3 filters. It was trained for 10 epochs on the training data. The SVM used a radial basis function (RBF) kernel and the KNN was tested with k=5 nearest neighbours.

For feature representation, color, texture and border descriptors were extracted. Color features included 11 Color Name (CN) linguistic labels. Texture was captured using SIFT and HOG descriptors. Border features like convexity, circularity, and irregularity index described lesion shape. The CNN automatically learned feature representations from the pixel data.

Classification results showed 85.5% accuracy with the CNN model, 71.8% with SVM, and 57.3% with 5-NN. By ensembling all models via majority voting, the overall accuracy improved to 88.4%. This demonstrates the power of combining multiple models to boost performance compared to individual models.

Zunair and Hamza [28] also employed a deep CNN architecture but focused specifically on acral melanoma, a rare subtype occurring on palms and soles. Their network had 9 layers and was trained on an augmented dataset of 640 images from the ISIC archive, resized to 124x124 pixels. Data augmentation techniques like rotation, flipping, blurring, and GAN-generated samples were utilized to balance the class distribution.

The architecture consisted of 3 convolutional layers for feature extraction, followed by 3 max pooling layers for downsampling. Two dropout layers were also added for regularization to reduce overfitting. The pixel-level input images were 224x224 in size. Without augmentation, the CNN achieved 92.7% testing accuracy on the ISIC dataset. The sensitivity was 95.26% and specificity 87.89%, showing reliable detection of both malignant and benign lesions.

Both studies performed vital preprocessing steps like hair removal and lesion segmentation to isolate the skin lesions. Daghrir et al. used morphological snakes for segmentation whereas Zunair and Hamza adopted gradient filtering.

Rashid et al. [29] proposed a MobileNetV2 transfer learning model, trained on 11,670 augmented ISIC 2020 images. They achieved 98.2% average accuracy for melanoma classification, with 98.3% recall and 98.1% F1-score. Extensive augmentation handled class imbalance.

Abdar et al. [30] evaluated uncertainty methods like Monte Carlo dropout on ISIC 2019. Their hybrid model combining dropout and ensembling attained 89% accuracy and 0.91 F1-score, showing improved diagnosis.

Preprocessing steps are vital for isolating lesions. Hasan et al. [31] used blackhat filtering for hair removal and morphological snakes for segmentation. Alwakid et al. [32] employed GANs for data augmentation and segmentation to extract region of interest.

For features, Alenezi et al. [33] extracted color, texture and shape descriptors. CNNs automatically learn hierarchical representations, e.g., MobileNetV2 and NASNet . Transfer learning avoids overfitting on small medical datasets.

Model design focuses on accuracy and efficiency. Lightweight SqueezeNet reduced parameters for edge devices . Spiking neural networks enable low-power hardware execution. Stacked pipelines improve segmentation, feature extraction and classification, but can increase overhead.

Kousis et al. [34] evaluated 11 CNNs on ISIC 2019, finding DenseNet169 achieved 92.25% accuracy. Mazoure et al. [35] developed a web-based classifier attaining 100% melanoma prediction probability. These represent state-of-the-art benchmarks, but real-world validation on diverse patients is still lacking.

Factors like skin type, hair, lesions and image artifacts continue to challenge existing methods. Large-scale annotated datasets with varied characteristics are lacking, limiting model generalization. Class imbalance also skews performance, requiring techniques like oversampling.

Key trends include personalized model optimization, uncertainty-aware prediction, edge-focused lightweight designs, and clinically-guided system development. With sufficient data and validation, deep learning holds immense potential to enhance melanoma screening, aiding early diagnosis and saving lives. But effective real-world translation remains a research priority.

In the domain of skin disease characterization, the XGBoost Classifier with VGG16 as an element extractor stands apart for its exceptional exactness, with a sensitivity of 95.26% and an specificity of 87.89% [ 22]. However, for those focusing on quick model preparation and expectation, LightGBM with VGG16 ends up being a more proficient decision, offering a sensitivity of 96.84% and a specificity of 82.89%. This compromise among exactness and effectiveness highlights the significance of choosing the right classifier in view of the particular necessities of the application [22]**.**

Factors like skin type, hair, lesions and image artifacts continue to challenge existing methods. Large-scale annotated datasets with varied characteristics are lacking, limiting model generalization. Class imbalance also skews performance, requiring techniques like oversampling.

Key trends include personalized model optimization, uncertainty-aware prediction, edge-focused lightweight designs, and clinically-guided system development. With sufficient data and validation, deep learning holds immense potential to enhance melanoma screening, aiding early diagnosis and saving lives. But effective real-world translation remains a research priority.

**References:**

[22] Melanoma classification using deep transfer learning Mounika, Yasaswi, Vandana, Sai Joshna Department of Computer Science and Engineering, Jain University, Bangalore, India

[23]Jinen Daghrir, Lotfi Tlig, Moez Bouchouicha, Mounir Sayadi. Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach. International Conference on Advanced Technologies for Signal and Image Processing, Sep 2020, Sfax, Tunisia. 10.1109/ATSIP49331.2020.9231544 . hal-03172718

[24]Skin Cancer Detection Using Combined Decision of Deep Learners AZHAR IMRAN 1 , ARSLAN NASIR2 , MUHAMMAD BILAL 3 , GUANGMIN SUN 1 , ABDULKAREEM ALZAHRANI 4 , (Member, IEEE), AND ABDULLAH ALMUHAIMEED 5

[25]Automatic Malignant and Benign Skin Cancer Classification Using a Hybrid Deep Learning Approach Atheer Bassel 1 , Amjed Basil Abdulkareem 2 , Zaid Abdi Alkareem Alyasseri 3,4,5,\* , Nor Samsiah Sani 2,\* and Husam Jasim Mohammed

[26]Skin Cancer Detection Using Combined Decision of Deep Learners AZHAR IMRAN 1 , ARSLAN NASIR2 , MUHAMMAD BILAL 3 , GUANGMIN SUN 1 , ABDULKAREEM ALZAHRANI 4 , (Member, IEEE), AND ABDULLAH ALMUHAIMEED 5

[27]Skin Cancer Detection using Ensemble Learning Dr. J. Avanija School of Computing Mohan Babu University Tirupati, India avans75@yahoo.co.in C. Chandra Mohan Reddy Dept. of CSE Sree Vidyanikethan Engg. College Tirupati, India chapireddychandu@gmail.com C. Sri Chandan Reddy Dept. of CSE Sree Vidyanikethan Engg. College Tirupati, India chandanreddy954@gmail.com D. Harshavardhan Reddy Dept. of CSE Sree Vidyanikethan Engg. College Tirupati, India harshavardhan.dasireddy@gmail.com Thumma Narasimhulu Dept.of CSE Sree Vidyanikethan Engg. College Tirupati, India thummanarasimha43@gmail.com Napa Venkata Hardhik Dept. of CSE Sree Vidyanikethan Engg. College Tirupati, India [nvhardhik21@gmail.com](mailto:nvhardhik21@gmail.com).

[28] Zunair, H., & Hamza, A. B. (2020). Melanoma detection using adversarial training and deep transfer learning. Physics in Medicine & Biology, 65(13), 135005

[29] Rashid, J.; Ishfaq, M.; Ali, G.; Saeed, M.R.; Hussain, M; Alkhalifah, T.; Alturise , F.; Samnd N. Skin Cancer Disease Detection Using Transfer Learning Technique. Appl. Sci. 2022, 12, 5714. <https://doi.org/10.3390/app12115714>

[30] Abdar, M.; Samami, M.; Mahmoodabad, S.D.; Doan, T.; Mazoure, B.; Hashemifesharaki, R.; Liu, L.; Khosravi, A.; Acharya, U.R.; Makarenkov, V.; et al. Uncertainty quantification in skin cancer classification using three-way decision-based Bayesian deep learning. Comput. Biol. Med. 2021, 135, 104418.

[31] Hasan, M.K.; Elahi, M.T.E.; Alam, M.A.; Jawad, M.T.; Martí, R. DermoExpert: Skin lesion classification using a hybrid convolutional neural network through segmentation, transfer learning, and augmentation. Inform. Med. Unlocked 2022, 28, 100819.

[32] Alwakid, G.; Gouda, W.; Humayun, M.; Sama, N.U. Melanoma Detection Using Deep Learning-Based Classifications. Healthcare 2022, 10, 2481.

[33] Alenezi, F.; Armghan, A.; Polat, K. Wavelet transform based deep residual neural network and ReLU based Extreme Learning Machine for skin lesion classification. Expert Syst. Appl. 2023, 213, 119064.

[34] Kousis, I.; Perikos, I.; Hatzilygeroudis, I.; Virvou, M. Deep Learning Methods for Accurate Skin Cancer Recognition and Mobile Application. Electronics 2022, 11, 1294.

[35] Mazoure, B.; Mazoure, A.; Bédard, J.; Makarenkov, V. DUNEScan: A web server for uncertainty estimation in skin cancer detection with deep neural networks. Sci. Rep. 2022, 12, 179.